



Ohio Commission on Minority Health

77 South High Street, 7th Floor, Columbus, Ohio
Phone: (614) 466-4000 FAX: (614) 752-9049

Evaluation Guidance for Health Grant Proposal Applicants

Prepared by the
RESEARCH AND EVALUATION ENHANCEMENT PROGRAM
EVALUATION PANEL

Betty Yung, Ph.D., Wright State University, Panel Chair
Marjorie Baker, Ph.D., Wright State University
Carla Clasen, M.P.H., Wright State University
Lucinda M. Deason, Ph.D., University of Akron
Rob Fischer, Ph.D., Case Western Reserve University
Peter Leahy, Ph.D., University of Akron
Fatima Perkins, M.N.O., United Way of Greater Cleveland
Manoj Sharma, Ph.D., University of Cincinnati

March, 2006



Ohio Commission on Minority Health

77 South High Street, 7th Floor, Columbus, Ohio
Phone: (614) 466-4000 FAX: (614) 752-9049

Evaluation Guidance for Health Grant Proposal Applicants

Table of Contents

	Page
I. BRIEF OVERVIEW OF OCMH’S VIEW OF EVALUATION.....	2
II. RATIONALE FOR EVALUATION.....	3
III. CULTURAL ISSUES IN EVALUATION.....	5
IV. HEALTH LITERACY ISSUES IN EVALUATION	10
V. EVALUATION LOGISTICS	
A. SELECTING AN EVALUATOR.....	11
B. EVALUATION BUDGET/CONTRACTING.....	11
C. INSTITUTIONAL REVIEW BOARD (IRB) PROCESS	12
VI. LOGIC MODEL AND PROGRAM OUTCOME MEASUREMENT	13
VII. THE OCMH’S SIX PRIORITY AREAS, EXAMPLES OF OUTCOMES AND OUTCOME INDICATORS	16
VIII. EVALUATION APPROACHES	
A. EVALUATION DESIGNS	18
B. INSTRUMENTATION, RELIABILITY AND VALIDITY	18
C. DATA COLLECTION METHODS	20
IX. ADDITIONAL RESOURCES.....	24
X. REFERENCES	28
APPENDICES	30
Appendix A. Bennett’s Model of Cultural Competence.....	31
Appendix B. Culturally Competent Methodologies	32
Appendix C. Outcome Indicators for Health Promotion/Disease Prevention	34
Appendix D. Glossary of Terms	43

PREFACE

Evaluation is an integral part of all planning efforts. When we conceptualize, write, or implement a program we should pay attention to evaluation very closely. This becomes more important for health programs as stakeholders, funders, and legislators demand more accountability, and as staff and administrators want to know more about the implementation and effectiveness of their programs.

The Ohio Commission on Minority Health (OCMH) funded a Research and Evaluation Enhancement Program (REEP) to bring together Ohio evaluation experts who have experience evaluating culturally diverse health research projects. The evaluation panel's role was to create a uniform, culturally competent, and scientifically sound evaluation system for the state's minority health projects. After an extensive state-wide search, a team of six evaluators who have expertise working with minorities was selected and the process for developing an evaluation model for projects serving minority populations was implemented. Several existing models such as the Centers for Disease Control and Prevention's evaluation model, Content, Input, Process, Product (CIPP) evaluation model, Kellogg Foundation's evaluation model, Predisposing, Reinforcing, and Enabling Constructs in Educational/Ecological Diagnosis and Evaluation & Policy, Regulatory, Organizational Constructs in Educational and Environmental Development (PRECEDE-PROCEED) evaluation model, and the United Way's evaluation model were carefully examined. Based on this review, appraisal of the evaluation results from the projects funded since 1987, and expertise of the evaluator panel, a culturally competent and practically feasible model that could be used in minority specific settings was developed. This model is presented in this document.

This model will help guide projects towards intended outcomes as well as streamline the data collection, data analysis, and reporting of evaluation results for the projects serving minority populations in the state of Ohio. The model can be replicated (in other states) for health projects that serve ethnically diverse minority populations.

Any questions about this document can be directed to the OCMH at (614) 466-4000.

This document is available online on the OCMH web site at <http://www.mih.ohio.gov>.

I. BRIEF OVERVIEW OF THE OHIO COMMISSION ON MINORITY HEALTH'S (OCMH)'S VIEW OF EVALUATION

The mission statement of the OCMH reads¹:

The Ohio Commission on Minority Health is dedicated to eliminating disparities in minority health through innovative strategies and financial opportunities, public health promotion, legislative action, public policy and systems change.

One of the ways that the OCMH seeks to deliver on this mission is by ensuring that funded projects systematically document and assess the delivery and effectiveness of their activities. In this sense, evaluation is a core component of each project. The OCMH seeks the following in the evaluation of the projects it funds:

- Evaluation methods should be appropriate to the intervention model being used in respect to scope and data collection.
- Each evaluation should include measures of both process outcomes (implementation) and behavioral/health status outcomes (changes in participants) relevant to the goals of the intervention.
- Evaluation is seen as a vehicle for program improvement (internal use) and program accountability (external use) and should provide information for the purposes of potential replication of the project.
- Evaluation findings from OCMH-funded projects will contribute to the existing knowledge base on what works in the field of minority health promotion.

¹ OCMH's Mission Statement is available at <http://mih.ohio.gov/Mission%20Statement.htm>

II. RATIONALE FOR EVALUATION

Why Measure Outcomes?

In growing numbers, service providers, government, other funders, and the public are calling for clearer evidence that the resources they expend actually produce benefits for people. Consumers of services and volunteers who provide services want to know that the programs to which they devote their time really make a difference. That is, they want better accountability for the use of resources provided. One clear and compelling answer to the question of "Why measure outcomes?" is: *To see if the OCMH funded programs in the priority areas of cancer prevention, cardiovascular disease and hypertension, diabetes, infant mortality, substance abuse, and violence really make a difference in the lives of African Americans, Latinos, Asian Americans, and Native Americans in Ohio.*

Although improved accountability has been a major force behind the move toward outcome measurement, there is an even more important reason: *To help programs improve the services they provide.* Outcome measurement provides a learning loop that feeds information back into programs on how well they are doing. It offers findings they can use to adapt and improve in order to become more effective.

This dividend doesn't take years to occur. It often starts appearing early in the process of setting up an outcome measurement system. The process of focusing on outcomes—on why the program is doing what it is doing and how it thinks participants will be better off—gives program managers and staff a clearer picture of the purpose of their efforts. Clarification alone frequently leads to more focused and productive service delivery.

Down the road, being able to demonstrate that their efforts are making a difference for people pays important dividends for programs. It can, for example, help programs:

- recruit and retain talented staff,
- enlist and motivate able volunteers,
- improve the conditions of participants,
- attract new participants,
- engage collaborators,
- garner support for innovative efforts,
- win designation as a model or demonstration site,
- retain or increase funding, and/or
- gain favorable public recognition.

Results of outcome measurement show not only where services are being effective for participants, but also where outcomes are not as expected. Program managers can use outcome data to:

- strengthen existing services,
- target effective services for expansion,
- identify staff and volunteer training needs,
- develop and justify budgets,
- prepare long-range plans, and
- focus board members' attention on programmatic issues.

To increase its internal efficiency, a program needs to track its inputs and outputs. To assess compliance with service delivery standards, a program needs to monitor activities and outputs. Additionally, agencies need to measure outcomes in order to improve their effectiveness in helping participants, assure potential participants and funders that their programs produce results, and show the general public that their services produce benefits that merit support.

These and other benefits of outcome measurement are not just theoretical. Scores of health professionals across the country attest to the difference it has made for their staff, volunteers, decision makers, financial situation, reputation, and, most important of all, for the public they serve.

III. CULTURAL ISSUES IN EVALUATION

Health Disparities: Understanding the Cultural Context for the Intervention

Prior to designing a program and evaluation that will help to reduce health disparities, the project director and evaluator must take into consideration the cultural context for the intervention and the participants' past experiences with the health care system. For instance, the Institute of Medicine (IOM) report titled, "Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care (2002) states:

Racial and ethnic minorities tend to receive a lower quality of healthcare than non-minorities, even when access-related factors, such as patient's insurance status and income, are controlled. The sources of these disparities are complex, are rooted in historic and contemporary inequities, and involve many participants at several levels, including health systems, their administrative and bureaucratic processes, utilization managers, healthcare professionals, and patients (p.1).

The IOM's report defined health disparities as "racial or ethnic differences in the quality of healthcare that are not due to access-related or clinical needs, patient preferences, and appropriateness of interventions." In the United States, racial and ethnic disparities are found in many sectors of society. As the IOM report notes:

African Americans, Hispanics, American Indians, and Pacific Islanders, as well as some Asian American subgroups are disproportionately represented in the lower socioeconomic ranks, in lower quality schools, and in poorer-paying jobs. These disparities can be traced to many factors, including historic patterns of legalized segregation and discrimination. Unfortunately, some discrimination remains. For example, audit studies of mortgage lending, housing, and employment practices using paired "testers" demonstrate persistent discrimination against African Americans and Hispanics. These studies illustrate that much of American social and economic life remains ordered by race and ethnicity, with minorities disadvantaged relative to whites. In addition, these findings suggest that minorities' experiences in the world outside of the healthcare practitioner's office are likely to affect the providers' and patients' perceptions in the healthcare setting (p.6).

The IOM's report indicate that the factors that play a major role in contributing to health disparities are at the healthcare system level (language barriers, time pressures on physicians, and geographic availability of healthcare institutions) and care process level (bias, stereotyping, and clinical uncertainty). The project director and evaluator must take these factors into consideration as they develop culturally competent programs and evaluations.

Culture and Cultural Competence

Culture encompasses race/ethnicity, gender, disability, religion, sexual orientation and other life-shaping factors. The phrases “cultural awareness” and “cultural sensitivity” have been used to describe the need for health care professionals to be responsive to culturally diverse clients. “Cultural competence” embodies going beyond awareness and sensitivity and actively incorporating cultural factors into the planning and delivery of programs and evaluation designs (Tripp-Reimer, 1997). Cultural competence has been defined as a “complex integration of knowledge, attitudes, and skills that enhances cross-cultural communication and appropriate/effective interactions” (Lenburg, et al., 1995). Cultural competence is viewed as a process, not an event, in which the provider of care and services continuously strives to achieve the ability to effectively work within the cultural context of an individual or community from a diverse cultural/ethnic background. To successfully provide services to all clients, it is important that the service providers understand each client’s frame of reference. Further, it is important not to consider cultural competency as a mere “skill building” activity. For it to be successful, it must, also, include policy and societal supports that are essential for cultural integration (Airhihenbuwa, 1995).

Cultural competence becomes important as members of our major ethnic groups are increasing in size, including African American, European American, Latino American, Native American, and Asian American and Pacific Islander populations. Additionally there has been an ideological shift in the nature of diversity in the U.S. While some people subscribe to the “melting pot” others focus on the preservation of distinct values and lifestyles, including expectations of health care (Deason and Wallace, 2005).

Given that the increasing size of the ethnically/racially diverse populations and various cultures, intervention programs and evaluation designs must become culturally competent. Nearly all work involves communication or the act of sharing information, yet we are rarely taught how to communicate effectively in a multicultural setting. It should be noted though that each member of a cultural group is a unique individual. Therefore, one should not attempt to place every member of an ethnic group into the same category. Program staff and evaluators must develop the skills necessary to obtain cultural information directly from participants, as each participant is different and may have experiences not shared by the cultural group to which she or he belongs. These skills prevent cultural knowledge from being applied stereotypically to all members of a cultural group. The Bennett model of cultural competence can help program staff and evaluators with becoming culturally competent (Bennett, 1993, see Appendix A).

Culturally Competent Evaluation

“Across the country, as health care providers work to deliver services in a culturally competent way, to address power imbalances and to legitimize cultural knowledge, leaders in health, evaluation and communities increasingly call for a parallel shift in the way we evaluate to be less exploitative, more respectful and inclusive of multicultural populations” (Endo, Joh, & Yu, 2003, p.1). There is very little literature on culturally competent evaluation. However, health and evaluation experts’ definition of culturally competent evaluation clusters into three areas: “1) the

characteristics of culturally competent evaluators (e.g., attitudes, skills and knowledge), 2) a culturally competent approach that includes and respects community voices in the evaluation, and 3) specific strategies in culturally competent evaluation design and implementation (Endo, Joh, & Yu, 2003, p.6-7). The root of cultural competency in evaluation starts with a genuine respect for the communities being studied and an openness to seek depth in understanding different cultural contexts, practices and paradigms of thinking. That is, culturally competent evaluators approach evaluations with a willingness to be creative and flexible in their evaluation designs to best capture different cultural contexts. Culturally competent evaluators also bring a heightened awareness of the power differentials that exist between an evaluator and those typically being evaluated.

Health and evaluation leaders also recognize that culturally competent evaluators also bring specific skill sets. These skill sets includes interpersonal skills such as the ability to build rapport across differences and gain the trust of community members as well as the ability to self-reflect and recognize the biases that one brings as an evaluator. In addition, the culturally competent evaluator should possess the ability to translate the jargon-laden field of evaluation to those not trained in this area and unfamiliar with the technical aspects of evaluation, or creatively modify traditional evaluation paradigms to be relevant within diverse cultural contexts. Much like cultural competency in the health care arena—it is dangerous to assume that cultural competency is based on knowledge about specific cultural groups. Rather, it is argued that culturally competent knowledge is rooted in a broader world perspective, often gained from experiences of living or working with multiple cultural groups. Furthermore, culturally competent evaluators should also have a fundamental understanding about historical and institutional oppression. This knowledge is critical for designing evaluations that integrate how historical and current social systems, institutions and societal norms contribute to disparities in different communities.

A Culturally Competent Approach

A culturally competent evaluation engages the community being evaluated with the evaluation itself. An evaluation approach that places an emphasis on community empowerment or community participation is necessary for a culturally competent evaluation. Experts identify participatory, empowerment, inclusive, responsive, utilization-focused, community based, or democratic deliberative evaluations as approaches that support cultural competence in evaluation.

Evaluation and health leaders stress that a collaborative approach with communities must be considered throughout the whole evaluation process, including developing the evaluation and research questions, designing the study, collecting data, analyzing data and reporting results. Collaborating with communities being studied gives the benefit of yielding more accurate and reliable data because an evaluator can double-check for bias in the evaluation design and data analysis.

Community participation and feedback are important for better and more accurate data because it also includes: 1) empowering groups being studied, 2) resulting in communities “owning” the

research process and the reasons for conducting the research; and 3) building the capacity of the participating organizations. For example, one evaluator conducted a project with a cultural group. This evaluator's design included training the community members in the value and utility of research and evaluation methods such as interviewing and conducting focus groups. Because of their increased skills, the community members ultimately ended up getting involved in developing new programs that were deployed in local units around the state. In contrast, in another project with different cultural groups, the evaluation was a disaster. None of the participants knew why they were completing a survey. Half of the people did not fill out the survey and the survey was not available in different languages. This was an example of a complete lack of understanding of the community. The lessons learned from this example is that without proper contextualization of the purpose of the survey and accessible data collection tools, well-intentioned evaluation efforts can easily backfire from distrusting community members.

Specific Strategies and Methodologies in Culturally Competent Evaluation

Finally, health and evaluation leaders also articulate *specific strategies* that they see as culturally competent. Most health and evaluation leaders point out that the evaluation questions and the conceptual framework should also guide the exact methods and strategies used to do an evaluation, with no one method or strategy ever ruled out. For example, one evaluator described different experiences he has had where one group of people believed that surveys were culturally irrelevant because their community had been “surveyed to death,” so they got their data by simply talking to people. Yet another community did not want oral interviews, so they used written surveys. Although they recognize the importance of the evaluation questions guiding culturally competent strategies, health and evaluation leaders offer these following strategies as examples of what they have found useful in their experiences.

Pre- Evaluation Work.

Health evaluators and leaders stress that taking the appropriate amount of time to plan is required to develop a successful evaluation plan. Taking the time to develop rapport over an extended period of time is necessary to gain credibility and legitimacy with groups of people. Evaluators need to take the proper time to do the necessary pre-evaluation implementation work in order to become familiar with the community context in which the evaluation will take place. One respondent gave an example of her pre-evaluation work in taking her community clients through developing a logic model (although she did not call it a logic model, which she believes is a very Western concept). Rather than go through a linear and sequential process to develop a model, she held a workshop to discuss the work the organization did and the reasons behind it. She pushed participants to discuss their underlying assumptions and their theory of change. Then she drafted a “conceptual picture” of the organization, which ended up looking like a traditional logic model, but was developed using a more conversational, collaborative style.

Post-Evaluation Work.

Finally, multicultural evaluation strategies extend beyond the design, data collection and analysis stages, to the post-evaluation work of reporting of findings. Recognizing that communicating the results of an evaluation is critical in how evaluation consumers use the results to make decisions and take their next steps, some health and evaluation leaders also include strategies for multicultural reporting in evaluation. A number state that there is too much emphasis on the written report format, which can alienate populations who are used to an oral tradition of communication, who are less inclined to read through long technical text or who would rather discuss findings and their meanings together as a group. One evaluator described a multimedia reporting format in which the evaluators incorporated poetry, collages, slides and music all with evaluation findings incorporated in them. Another evaluator describes a “pre-pre-pre-reporting” phase during a preliminary analysis of findings, in which the evaluator and those being evaluated share ideas about the findings. This “siphoning” of data and data analysis reflect 1) the evaluator’s sense of accountability to the community; and 2) an ethical and sincere commitment to getting feedback from the community to see if the evaluator got her/his interpretations and facts right.

Overall, health and evaluation leaders believe that — despite advances in institutionalizing culturally competent service delivery — widespread adoption of multicultural evaluation approaches in this field to date has been less pervasive than in other fields such as education or anthropology. They express, however, a sense of energy around building upon the multicultural evaluation movement within the broader evaluation field. They also articulate a tremendous opportunity to advance multicultural evaluation within the field of health. According to health and evaluation leaders, the growing presence of different communities of color is prompting a broader “paradigm shift” in how the field is thinking about multiculturalism, health care delivery, and subsequently research and evaluation. Given the

My [evaluation] model is empowerment and capacity building. I may do some work, but I’m always doing work so that people a) understand it, and b) own up to the process and the data, and I make sure that whatever I do I siphon it back to them so that it helps them. —Chwee Lye Chng, health researcher and evaluator

rapidly growing diversity of health care constituencies and influential research from the Institute of Medicine, the U.S. Surgeon General, the Centers for Disease Control and Prevention and others that continue to document persistent disparities in health status across multicultural populations culturally competent evaluation methods and research are needed. From the bottom-up, community-level advocacy and increased expression around issues of equitable access to care is also forcing research and evaluation communities to consider approaches to and basic assumptions about research and evaluation. Finally, some maintain that the very nature of health requires increased focus on multicultural evaluation. The unique integration of health and culture — from culturally determined basic conceptions of “health” or “illness” to cultural stigma associated with HIV/AIDS, depression or sexually transmitted disease — is demanding that professionals collect data and conduct meaningful analysis in culturally competent ways in order to ultimately impact the improved health status of multicultural communities.

Methodologies for Culturally Competent Evaluation

Various qualitative and quantitative methodologies facilitate conducting culturally competent evaluations. Some of these methodologies include: mixed method approaches, triangulation, reflexive staffing, and engagement of community members (participatory research). Specific details about these methodologies can be found in Appendix B.

IV. HEALTH LITERACY ISSUES IN EVALUATION

Health literacy is a constellation of skills, including the ability to perform basic reading and numerical tasks required to function in the health care environment. In various settings, patients with adequate health literacy can read, understand and act on health care information (Bresolin, 1999). More specifically, health literacy is the degree to which individuals have the capacity to

- obtain,
- process, and
- understand basic health information and services needed to make appropriate health decisions (National Library of Medicine, 2000).

Low literacy may impair functioning in the health care environment or participation in health promotion programs, affect patient-health provider communication dynamics, and inadvertently lead to substandard medical care or poorer compliance with health promotion program activities and suggestions. It is, also, associated with poor understanding of written or spoken health advice and adverse health outcomes. Although low health literacy is more prevalent among people who completed fewer years of education, persons of certain ethnic/racial groups, the elderly, and individuals with low cognitive ability, the Institute of Medicine (2004) report noted that people of all ages, races, incomes and education levels are challenged by low health literacy (Institute of Medicine, 2004). Some other factors associated with low literacy include living in the South or Northeast (rather than the West and Midwest), being female, being incarcerated, and being classified as poor.

Instruments have been developed for measuring literacy in the health care setting including the Wide Range Achievement Test (WRAT), Rapid Estimate of Adult Literacy in Medicine (REALM), and the Test of Functional Health Literacy in Adults (TOFHLA). These instruments have focused on the ability to read and use numbers. Studies have found a positive and significant relationship between literacy levels and participants' knowledge of health issues (Arnold et al., 2001; Williams, Baker, Honig, et al., 1998; and Williams, Baker, Parker, et al., 1998). Interventions and evaluations designed for the low health literacy population must take into account the participants' ability to understand, process and act on health information.

Most studies addressing health outcomes focused on improvements in knowledge (Berkman, et al., 2004). In the low health literacy population, some interventions that have helped to improve knowledge include using a videotaped educational tool (Murphy, 2000), utilizing illustrated materials (Michielutte et al., 1992), and designing pamphlets for easy readability (Davis et al., 1998). It is crucial that project directors and evaluators take into consideration the health literacy level of the target population and develop appropriate interventions and tools. It often helps to

have representatives from the target audience who are familiar with the health literacy level of their community in developing and pilot testing evaluation instruments.

V. EVALUATION LOGISTICS

Just as the development and implementation of health promotion programs and interventions involve a considerable amount of planning and coordination, so too do the evaluation of these efforts. This section provides information on selecting of an evaluator for your project, budgeting for and contracting with this individual or institution, and planning for the protection of human subjects.

A. SELECTING AN EVALUATOR

When selecting an evaluator for your OCMH project, the following criteria may be useful to consider:

- **Credentials/Reputation-** To what extent has the individual evaluated programs, particularly minority focused programs? Is the evaluator affiliated with an academic institution?
- **Education-** To what extent the individual trained in evaluation? Does the evaluator have a certificate, bachelor, master or doctoral degree related to evaluation?
- **Experience-** To what extent does the individual have formal or informal experience with evaluation in public health settings? Does the evaluator specialize in one type of evaluation? Is the evaluator able to help develop necessary forms and processes to evaluate the program? Can the evaluator provide samples of evaluation reports that he or she has completed?
- **Cultural Competency/Sensitivity-** To what extent does the evaluator have experience working with the target population? Has the evaluator participated in cultural competency trainings or received a certificate in cultural competency.
- **Integrity-** To what extent is the organization familiar with the evaluator? Does the evaluator present any conflicts of interest with evaluating the project?
- **Communication Skills-** To what extent is the evaluator able to explain technical concepts in understandable language and demonstrate clear verbal and written expression?
- **Availability-** To what extent is the evaluator available to meet with the project stakeholders, adhere to project timelines and be flexible if timelines need to be modified?
- **Cost-** Are the proposed cost for the evaluation reasonable for the required tasks?
- **Contract/Scope of Service-** Is the evaluator willing to design a scope of service/contract or agreement outlining his/her responsibilities along with timelines?

There is additional information that provides guidance on selecting an evaluator in the resources section of this manual.

B. EVALUATION BUDGET/CONTRACT

As with all parts of a public health project sufficient resources should be set aside for the evaluation of the initiative. Ideally an evaluation partner will be identified prior to the submission of an application to OCMH. In this case the evaluator can assist you in developing a realistic budget for the evaluation of the project, based on the scope of work, methods, time required, etc. As a rule of thumb, projects could expect to allocate resources in the range of 6-12% of the cost of delivering the program to the evaluation. Thus, for a \$250,000 project award, the evaluation could involve costs of \$15,000 to \$30,000.

If the project uses an individual consultant as an evaluator, the evaluation budget should include a description of effort and compensation rate. This could be the number of days or hours and the rate that the evaluator will charge or it could be a flat fee for the completion of specific deliverables. If the project expects to use an institutional partner (i.e., university or research firm), the budget will be very similar, except these institutions will also include an indirect or overhead charge on the evaluation contract. Usually these indirect rates are negotiable but could range from 20-50% of the direct costs of the research, depending on the institution. It is helpful for community-based agencies to develop written policies about the level of indirect costs they will support in contracts with outside organizations. Such policies may be helpful in the negotiation process.

Projects should expect to develop a written agreement/contract with the evaluator they have chosen. The agreement should clearly identify the tasks to be completed, the timeline for completion of tasks, and the deliverables that the evaluator is to produce. In respect to formal agreements, projects are encouraged to consider the following recommendations:

- Include the evaluation design in any formal agreements.
- Do not expect participation in the evaluation by persons who have not previously agreed to do so.
- Do not act unilaterally in a matter where it has been agreed that evaluator/client collaboration would be required for decisions.
- Do not change the design without amending formal agreements.²

C. INSTITUTION REVIEW BOARD (IRB) PROCESS

If you are working with an academic institution, your evaluator will likely be involved in a review process with the college or university's Institutional Review Board (IRB). The IRB is a committee established to protect the rights and welfare of the human subjects involved in research activities such as program evaluation. IRB committees review proposed research plans to ensure that researchers educate potential participants about the purposes, risks, and benefits of the research, inform them of the voluntary nature of their participation, their right to withdraw at

² Joint Committee on Standards for Educational Evaluation (1994).

any time and their right to receive information about the research results. IRB committees also make sure that written consent has been received from participants as well as ensuring a plan is in place to keep participant information confidential. It is important to keep in mind that the IRB process generally takes several weeks to complete and may add additional time to the start-up of the project.

VI. LOGIC MODEL AND OUTCOME MEASUREMENT

The OCMH plans to fund projects that develop logic models that clearly show the link between program inputs, activities, outputs, outcomes, and outcomes indicators. These types of programs help the OCMH to accomplish parts of its mission. The logic model is a planning tool used to clarify and illustrate what your project intends to do and what it hopes to accomplish and impact. More specifically, a logic model:

- summarizes key program elements;
- explains rationale behind program activities;
- shows the links between program inputs, activities, outputs, outcomes, and outcome indicators;
- clarifies intended outcomes; and
- provides a communication tool.

The logic model can be considered a map that you develop to clarify and communicate what your project intends to do and its presumed impact. It consists of program inputs, activities, outputs, outcomes, and outcome indicators. These components are linked together to form the building blocks of a logic model. A logic model template is shown in Figure 1. Figure 2 provides an example of a well constructed logic model for a prevention and early detection of cancer in youth and adults program. This cancer prevention program will consist of providing participants with information and useful materials related to cancer that will: 1) enable them to make informed decisions about prevention, early detection, and treatment, 2) encourage participants to follow recommended screening guidelines for prevention and early detection of cancer, and 3) provide participants with examinations for skin cancer and information about risk reduction and referrals for treatment if needed. The program inputs, activities, outputs, outcomes, and outcome indicators are adequately described and linked together (see Figure 2).

FIGURE 1: GENERIC LOGIC MODEL

Inputs	Activities	Outputs	Outcomes	Outcome Indicator(s)
<p>What is used to deliver services</p> <p>Resources dedicated to or consumed by the program: ✓ Staff & staff time ✓ Volunteers & volunteers' time ✓ Facilities ✓ Equipment & Supplies</p> <p>Constraints on the program: ✓ Laws ✓ Regulations ✓ Funder's requirements</p>	<p>Things done to, for or with the target population intended to lead to the desired change</p> <p>What the program does with inputs to fulfill its mission: ✓ Conduct a breast cancer screening program ✓ Educate the participants about benefits of physical activity ✓ Counsel pregnant women about nutrition ✓ Educate the participants about weight loss</p>	<p>Number of people served, hours of instruction provided and received</p> <p>The direct products of program activities: ✓ Number of classes taught ✓ Number of educational materials distributed ✓ Hours of service delivered ✓ Number of participants served</p>	<p>Benefits as a result of the activities conducted</p> <p>Benefits or changes for participants during or after program activities: <u>Initial</u> ✓ New knowledge ✓ Increased skills ✓ Changed attitudes or values</p> <p><u>Intermediate-term</u> ✓ Modified behavior</p> <p><u>Long-term</u> ✓ Improved condition ✓ Altered status</p>	<p>What we track and use to measure the benefits to program participants</p> <p>✓ Number and percent of participants that learned how to lose weight (pre- and post-test surveys will be administered)</p> <p>✓ Number and percent of participants that lost weight (a scale will be used to measure participants weight before and after the weight loss class)</p>

**FIGURE 2. PREVENTION & EARLY DETECTION OF CANCER FOR YOUTH & ADULT PROGRAM
LOGIC MODEL**

I. Program Inputs

- A. Volunteers
- B. •Executive Director,
•Cancer Control Specialists,
•Cancer Control Director,
•Administrative Assistant,
•Generalists,
•Nurse
- C. Knowledge of Issues
- D. Community Partners/
Collaborators
- E. Medical equipment & supplies through community partners (e.g. Mobile Mammogram units)
- F. Office space & equipment,
- G. Funding Source Funding (\$125,000)

II. Program Activities

- A. Disease specific information provided, support, and follow-up (Related inputs are: I.A, I.B, I.C, I.D, I.E, I.F, & I.G)
- B. Mammography screening program provided for women (Related inputs are: I.A, I.B, I.C, I.D, I.E, I.F, & I.G)
- C. Adult intervention programs: “Let’s Talk About It for African Americans, Man-to-Man Prostate education support program, skin cancer education and screening (Related inputs are: I.A, I.B, I.C, I.D, I.E, I.F, I.F, & I.G)
- D. Body & Soul: A celebration of healthy eating (Related inputs are: I.A, I.B, I.C, I.D, I.E, I.F, & I.G)
- E. Stay Tobacco-Free Athlete Mentor Program for youth (Related inputs are: I.A, I.B, I.C, I.D, I.E, I.F, & I.G)
- F. Workplace Wellness: Menu of training choices, Active for Life, 5-a-day freshstart, tobacco cessation, meeting well, We Care Newsletter, and a website (Relate inputs are: I.A, I.B, I.C, I.D, I.E, I.F, & I.G)

III. Program Outputs (per year)

- A. 2,000 calls/year received outside of the call center & direct staff calls (Related activities are: I.A, I.B, I.C, I.D, I.E, & I.F)
- B. 800 women screened/year (Related activities are: I.A, I.B, & I.C)
- C. 22 one-hour sessions year reach an estimated 500 clients (Related activities are: I.A, I.C, I.D, I.E, I.F)
- D. Quarterly meetings are conducted each year reaching an estimated 200 (Related activities are: I.A, I.C, I.B, I.D, I.E, & I.F)
- E. 1,300 screened during May (Related activities are: I.A, I.B, & I.C)
- F. 10 African American religious institutions reach an estimated 3,000 members (Related activity is: I.A)
- G. 5 high school mentor reach 1,000 middle school students (Related activity is I.E)
- H. Reach 4,000 employers/ employees (Related activity is: I.F)

IV. Program Outcomes

- INITIAL OUTCOMES**
- A. Participants receive cancer information that enables them to make informed decisions about cancer prevention, early detection and treatment (Related outputs are: I.A, I.B, I.C, I.D, I.E I.F, I.G, & I.H)
 - B. Participants will be examined for skin cancer, receive information on risk reduction and referrals for tests/treatment if needed (Related outputs are: I.A, I.E, I.F, & I.H)
 - C. At-risk women will have access to mammograms and receive referral for further tests and assistance as needed (Related outputs are: I.A, I.B, I.E, I.F, & I.H)

V. Outcome Indicators

- A. # of callers or consumers who receive information and/or referrals (Related outcomes are: I.A, I.B, & I.C)
- B. # of individuals who were examined and/or referred from the skin screening (Related outcomes are: I.A & I.B)
- C. # of women screened who are referred for mammograms and/or additional medical consultations (Related outcomes are: I.A & I.C)

VII. THE OCMH'S SIX PRIORITY AREAS, EXAMPLES OF OUTCOMES AND OUTCOME INDICATORS

The OCMH seeks to fund programs in six priority areas, including: 1) cancer education/behavior change and prevention, 2) cardiovascular disease/hypertension education/behavior change, 3) diabetes education/behavior change, 4) infant mortality education/behavior change, 5) substance abuse education/behavior change, and 6) violence education/behavior change. Outcomes are specific statements about the desired changes in the lives of the participants. Outcome levels are initial (very close in time to the program intervention), intermediate (farther out in time, perhaps several months after the intervention), and long-term (even farther out in time, perhaps more than a year after the program). In addition, short-term outcomes involve changes in learning, awareness, knowledge, attitudes, skills, opinions, aspirations and motivation. Intermediate outcomes include changes in action, behaviors, practices, decisions, policies and social action. Lastly, long-term outcomes involve changes in the broader societal health, civic and environmental conditions. Examples of short-term, intermediate-term and long-term outcomes in each of the priority areas can be found in Appendix D.

Table 1 provides examples of some outcomes and outcome indicators. The examples in Table 1 are **illustrative examples only**. Programs need to identify their own outcomes and indicators, matched to and based on their own experiences, missions and input of their staff, volunteers, participants and others.

TABLE 1. EXAMPLES OF OUTCOMES AND OUTCOME INDICATORS

Type of Program	Outcome	Outcome Indicator(s)
Cancer Education/Behavior Change	<p>Increased knowledge about cancer prevention strategies</p> <p>Increased number of people receiving preventive screenings for cancer</p>	<ul style="list-style-type: none"> • 80% of program participants will demonstrate increased knowledge about how healthy diet choices reduce the risk of cancer based on a self-report test. • 30% increase in number of women over 40 in the target population who have received a mammogram in the past two years based on self-report.
Cardiovascular Disease/Hypertension Education/Behavior Change	<p>Increased ability to manage stress by program participants</p> <p>Increased levels of physical activity by program participants</p>	<ul style="list-style-type: none"> • 90% of program participants will be able to describe at least one technique for stress reduction. • 50% of program participants will report using stress reduction techniques • 50% increase in number of adults in the target population who report getting at least 20 minutes of physical activity at least 3 times a week.
Diabetes Education/Behavior Change	<p>Increased knowledge about managing diabetes</p> <p>Decreased number of complications of diabetes</p>	<ul style="list-style-type: none"> • 80% of participants will demonstrate increased knowledge of healthy dietary choices. • 90% of diabetic patients in the target population will have a documented foot exam in the past year.
Infant Mortality Education/Behavior Change	<p>Increased frequency of prenatal care visits</p> <p>Increased use of prenatal vitamins</p>	<ul style="list-style-type: none"> • 60% increase in number of participants completing all recommended prenatal visits prior to delivery • 90% of participants will report daily use of prenatal vitamins
Substance Abuse Education/Behavior Change	<p>Increased knowledge about the effects of alcohol on the body</p> <p>Decreased reported use of marijuana by adolescents aged 14-17.</p>	<ul style="list-style-type: none"> • 80% of participants will demonstrate an increase in knowledge of alcohol's effects on the brain, liver, and stomach. • 80% of students who report no use of marijuana at the beginning of the program will continue to report no use at the end of the program period.
Violence Education/Behavior Change	Decreased incidence of fighting at school	<ul style="list-style-type: none"> • 20% fewer disciplinary actions for fighting at school by the end of the program, compared with the number of disciplinary actions in the year before the program.

VIII. EVALUATION APPROACHES

A. EVALUATION DESIGNS

Evaluation Designs

Evaluation “design” is a term used by evaluators to describe a scientific strategy to systematically measure program effect and to attribute that effect to a program’s intervention. A large number of evaluation designs are available for use by grantees.³ Evaluators recognize that some designs are much “stronger” than others, however. Strength is measured by the design’s ability to enable the evaluator to conclude that a program effect is due to the program and not to something else. The “something else” refers to all the other possible explanations that typically exist for a program’s effect, usually referred to as the “threats to the internal validity” of a program.⁴ All OCMH grantees are required to identify and discuss an evaluation design in their evaluation proposal.

While the “gold standard” in evaluation design in general is the randomized comparison group design, the OCMH understands that this design is unlikely to have applicability in a local community setting. Therefore, the OCMH recommends at a minimum that a one-group pre-test/post-test design be used with outcome measures collected from all program participants using valid and reliable instruments. The pre-test/post-test design can be strengthened considerably by adding a similar comparison group for the data collection who will not be involved during the programming phase. No matter what design you choose, however, baseline (pre-implementation) data must be collected on program participants before the start of the program to establish a benchmark to compare change against. You will then be expected to collect the same information from participants after the program is complete.

Choosing Your Evaluation Design

Evaluation designs must be selected to fit a particular program’s situation. There is no “one size fits all” evaluation design in most cases. Grantees are **strongly encouraged** to discuss their evaluation design with their evaluation consultant at the proposal development stage. It is easier to “fit” an evaluation design to a program at the program planning stage than it is after implementation. Your evaluation consultant will guide you on the proper selection of an appropriate evaluation design.

B. INSTRUMENTATION, RELIABILITY AND VALIDITY

An instrument is what we use to obtain information on the knowledge, skills, attitudes and behavior of participants in a health promotion project. The term instrument is often used interchangeably with the terms tool, survey, measure, measurement, and questionnaire. It may be also referred to as a scale, or a cluster of items that measure a single concept; for example, an

³ You may have heard of some of these designs referred to by such terms as the one group pre-test/posttest design, the randomized comparison group design, or simply as a quasi-experimental design.

⁴ See glossary of terms for a definition of internal validity.

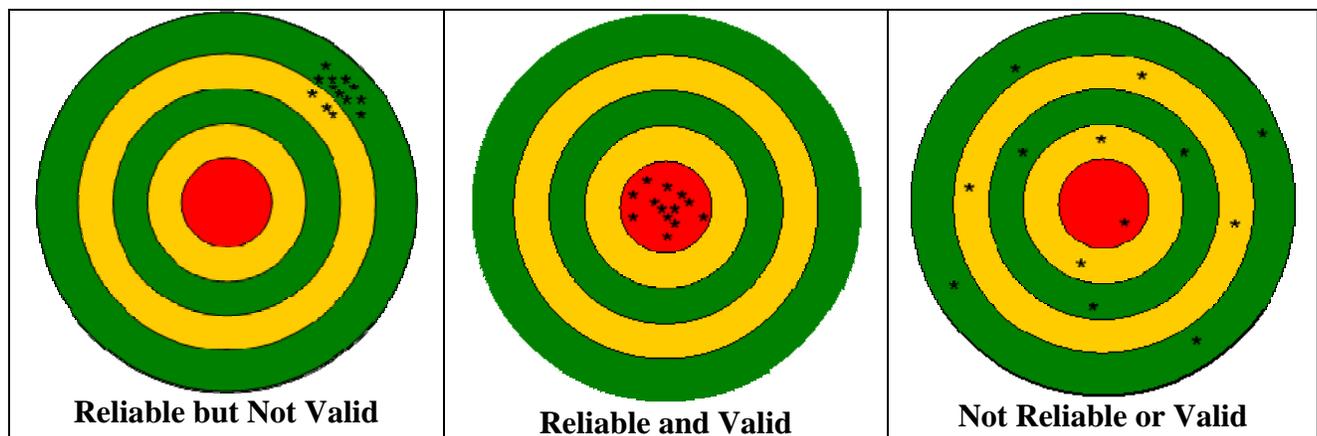
instrument measuring diabetes-related behavior might have a scale that measures “knowledge about diabetes” and another that measures “actions taken to manage diabetes.”

One way in which the quality of an instrument can be verified is by knowing that it is reliable and valid.

Reliability. Reliability refers to the consistency of an instrument to yield the same results when the instrument is used to measure the same thing time and time again. An instrument is said to be reliable when it consistently measures each time it is used. A reliable instrument is like a reliable watch—its accurate keeping of time is not affected by the race, weight, height, or income of the person wearing it nor by the month of the year, day of the week, or temperature of the room in which it is worn. If an item on a survey is unclear, vague, or subject to various interpretations, then it might be understood differently by the person responding and the responses to the survey questions would therefore not be reliable.

Validity. Validity refers to the degree to which an instrument measures what it was designed to measure. For example, if you are trying to measure the impact of a diabetes education program focusing on healthy nutritional choices but your instrument only asks questions about attitudes toward food, it would not be a valid measure of the effects of your health education intervention.

One way to understand the concepts of reliability and validity and their importance to evaluation research is to envision an archer and a target.⁵ In this context, an evaluation instrument is like an arrow shot at a target and the bulls’ eye on the target is what we are trying to measure. Over repeated attempts to connect with the target we will be able to observe the level of consistency (reliability) and accuracy (validity) of our “arrow.” The center panel below reflects the experience of using an instrument that is both reliable and valid. To the left we can see the effects of using an instrument that is reliable but is off target for what we are seeking to learn. To the right, our instrument is both unreliable and not valid, resulting in scattered shots that do not hit the bulls’ eye.



⁵ Graphic adapted from <http://www.georgetown.edu/departments/psychology/researchmethods/researchanddesign/validityandreliability.htm>

The reliability and validity of an instrument has a direct impact on the confidence that can be placed in the evaluation findings. This, in turn, significantly impacts the usefulness of the evaluation findings in planning future programs and in anticipating future outcomes.

Many of the measures of the effects of a health promotion are clinical indicators (e.g., blood pressure readings) or behaviors directly observed (e.g., participants exercising in a fitness program). While they do of course require documentation, they obviously do not involve instrumentation that needs validity and reliability established. However, many intermediate outcome indicators (e.g., attitudes toward participating in an exercise program or self-reported behaviors to lower blood pressure) do require reliable and valid measures to assess changes in participant attitudes or behaviors. **Whenever possible, it is highly desirable to use standardized instruments that have already been established as having sound reliability and validity.** Use of these instruments has the advantage of yielding evaluation findings that can be presented with greater confidence and credibility.

Unfortunately, there is still a scarcity of cross-culturally validated instruments in health research as well as instruments available in a variety of languages. Sample standardized instruments that are valid and reliable and that have been used frequently with minority populations are provided for each of the six (6) outcome indicators and may be found in a separate report on the OCMH's website entitled *Standardized Instrument Inventory for Measuring Racial/Ethnic Health Promotion Program Outcomes*.

The lack of availability of instruments in some areas may necessitate the development of a new instrument. In such cases it is important for the agency and the evaluator to work together to assess the new instrument's reliability and validity (sometimes referred to as psychometric properties). To learn more about this process, readers are referred to: DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd ed.). Thousand Oaks, CA: Sage Publications.

C. DATA COLLECTION METHODS⁶

A key aspect of evaluation involves the selection of appropriate data collection methods. Table 2 (page 23) presents information about six types of data collection methods that may be relevant in specific program environments. The table shows each method's primary purpose and the corresponding advantages and challenges.

Using Single Versus Multiple Methods

No one data collection method is ideal for every situation. For this reason, it is preferable to use multiple methods whenever possible. Using multiple methods to assess the same outcomes (e.g., using surveys and document reviews to assess program management) provides a richer, more detailed picture. It also illuminates inconsistencies between methods and reduces the chance of bias caused by a particular method.

⁶ Section extracted from: Bouffard, S., & Little, P. M. D. (2004).

Quantitative and Qualitative Evaluation Methods

Quantitative evaluation methods are required in most evaluation situations. Quantitative evaluation attempts to use systematic methods to quantify or measure program outputs and outcomes. Thus, a quantitative evaluation of a typical program will count the number of program sessions offered to program participants, count the number of clients who attended the program sessions or ask participants for their evaluation of service delivery on a standardized satisfaction survey at the completion of the program. Similarly, a quantitative evaluation attempts to measure outcomes such as changes in the skills, attitudes or behaviors of program participants using reliable and valid instruments. Quantitative evaluation can be strengthened considerably by also adding qualitative approaches.

In qualitative evaluation the meaning of any situation is perceived from the participant's viewpoint, and the researcher is a mere observer. Qualitative evaluation attempts to be holistic in its orientation to construct and interpret reality from multiple sources.. Qualitative designs are dynamic and based on inductive reasoning. Some commonly used methods that employ qualitative designs for conducting health program evaluations include: ethnographies, phenomenologies, interviews, case studies and participatory research studies (Sharma, 2004). The sampling techniques that are often used for qualitative evaluation include purposive and convenience sampling. Interviews, focus groups and observations are common methods for collecting data. Content analysis is then used to analyze the results from the data collected. Issues of trustworthiness (adherence to procedures and rigor), credibility (consensus), coherence (extent to which the final evaluation makes sense), transferability (ideas about similarity of evaluation in other settings), dependability (similarity with other evaluations), and confirmability (extent to which neutrality of evaluation can be established) are important considerations. For an example of participatory research see a study conducted by Sharma & Deepak (2001). The data collection methods mentioned above are summarized in Table 2 on page 23.

Selecting Data Sources

Equally important in selecting a data collection method is selecting a data source or information provider. Data can be collected from youth, families, staff, funders, educators, and other stakeholders. Certain data sources lend themselves more easily to certain data collection methods; for example, to assess parent satisfaction, surveys are popular because they can be mailed to hard-to-reach parents. As with data collection methods, it is recommended that evaluations use several data sources. However, programs may already have some of the evaluation information they need. Conducting an inventory of what programs already have and what is already available in the community can reduce the data burden as well as data collection costs.

Selecting a Sample of Individuals

The choice of data collection methods is also affected by the sample to be studied. Some methods are well suited to collecting data from all participants (e.g., surveys), while others are better suited for a smaller group that represents the diversity of all participants (e.g., focus groups). The choice of sample is in turn affected by the size of the program to be evaluated. In general, including all participants produces the most reliable results but may limit the type and amount of data collected because of the cost implications of doing so.

Collecting Data Before and After Program Participation and Implementation

Using the same data collection method to gather information before the start of the program and after its completion (also known as a *pre-/post-test design*) provides the opportunity to determine whether some characteristic changed during the course of the program. This can *suggest* that the program played a role in effecting the change; however, unless a program uses random assignment as part of its pre/posttest design, this method cannot establish that the program *caused* the change, because other unmeasured factors may have been responsible.

Cost Considerations

Selecting data collection methods has cost implications. While conducting observations and reviewing program documents can be done with little additional funding, designing and administering a survey and analyzing its results may require the assistance of an outside evaluator. At a minimum, it will require additional staff time for training. Similarly, using standardized tests and assessments will likely necessitate external expertise to analyze results. An additional cost consideration is the use of management information systems (MIS) that can be used to record and store data collected from numerous methods. Like most other components of evaluation, it is important to start small and build data collection tools over time, as appropriate to the program's evolving evaluation needs.

Table 2: Data Collection Methods

Method	Purposes	Advantages	Challenges
Surveys and Questionnaires	Collected by evaluators to gather specific information from participants, families, staff and administrators, community members, and other stakeholders. Usually administered on paper, in a structured or semi-structured format. Respondents often choose from among a set of <i>forced-choice</i> , or provided, responses. Can be administered in person, by mail, over the phone, or via email/Internet.	<ul style="list-style-type: none"> <input type="checkbox"/> Less time consuming and expensive to administer. <input type="checkbox"/> Can be administered to large groups of individuals. <input type="checkbox"/> Effective for assessing program satisfaction. 	<ul style="list-style-type: none"> <input type="checkbox"/> Data entry and analysis can be time consuming. <input type="checkbox"/> May be difficult to receive completed surveys from stakeholders. Incentives can be offered to increase response rate.
Interviews and Focus Groups	Most often used to gather detailed, qualitative descriptions of how programs operate and how stakeholders perceive them. Interviews are conducted one-on-one, while focus groups are conducted in small groups. Usually conducted with targeted samples of stakeholders. Can be conducted in person or by phone. Questions are generally open-ended and responses documented in detailed notes or transcription.	<ul style="list-style-type: none"> <input type="checkbox"/> Provide rich data that paint a broad picture. <input type="checkbox"/> May highlight information useful for interpreting quantitative data. <input type="checkbox"/> Small focus groups may increase the comfort level of participants. 	<ul style="list-style-type: none"> <input type="checkbox"/> Interviews and focus groups may intimidate some participants. <input type="checkbox"/> Documentation and analysis can be time consuming and may require the help of someone versed in qualitative analysis.
Observations	Generally an unobtrusive method for gathering information about how the program operates. Usually conducted by external evaluators and are often used to supplement information gathered through other methods. Can be highly structured or unstructured. More reliable if conducted over time to lessen the chance that observations are atypical.	<ul style="list-style-type: none"> <input type="checkbox"/> Provide highly detailed information from an external perspective on what occurs. <input type="checkbox"/> Trained observers may provide less biased descriptions. 	<ul style="list-style-type: none"> <input type="checkbox"/> Labor intensive and expensive. <input type="checkbox"/> Observers must be trained and be consistent. <input type="checkbox"/> Observations may not represent practices over time.
Tests and Assessments	Developed or used specifically for the program evaluation to quantify characteristics of the program, participants or outcomes. Examples include physical exams and psychological tests. May be standardized or created by program evaluators for the specific program.	<ul style="list-style-type: none"> <input type="checkbox"/> Often more valid and reliable than perceptions or opinions. <input type="checkbox"/> Comparing pre-/post-test scores is a good method for assessing change. 	<ul style="list-style-type: none"> <input type="checkbox"/> Can be costly and time intensive. <input type="checkbox"/> May require scoring by an external source.
Record or Document Reviews	Analyze existing program records and other documents not gathered or developed specifically for the evaluation. Examples include recruitment and attendance records, budget, staff records and annual reports. Particularly useful for documenting implementation.	<ul style="list-style-type: none"> <input type="checkbox"/> Records are tailored to programs. <input type="checkbox"/> Saves on evaluation time/cost. <input type="checkbox"/> May elicit a high degree of accuracy if records kept. 	<ul style="list-style-type: none"> <input type="checkbox"/> May not be available or applicable for some indicators. <input type="checkbox"/> May be incomplete due to staff time constraints.
Secondary Sources and Data Reviews	Use existing documents or data that were originally collected for other purposes, but which are useful for the evaluation. Examples include achievement data, standardized test scores, court records and community demographic data.	<ul style="list-style-type: none"> <input type="checkbox"/> May be less biased than perceptions or opinions. <input type="checkbox"/> Can save administrative time and costs. 	<ul style="list-style-type: none"> <input type="checkbox"/> Obtaining records often requires special permission. <input type="checkbox"/> Some national datasets charge access fees.

Summarized from: Bouffard, S., & Little, P. M. D. (2004). *Detangling Data Collection: Methods for Gathering Data*. Harvard Family Research Project, Harvard Graduate School of Education.

IX. ADDITIONAL RESOURCES

Cultural Competence

- Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.). *Education for the intercultural experience*. Yarmouth, ME: Intercultural Press.
- Campinha-Bacote, J. (1994). The process of cultural competence in health care: A culturally competent model of care. Cincinnati, OH: Transcultural C.A.RE. Associates.
- Chrisman, N.J. (1991). "Cultural systems." In S. Baird, R. McCorkle, and M. Grant (Eds.). *Cancer nursing: A comprehensive textbook*. Philadelphia PA: W.B. Sanders & Co.
- Deason, L. M. & Wallace, S. L. (2005), Does a shared culture between nursing home staff and residents influence quality of life?", *International Journal of Diversity in Organisations, Communities, and Nations*, Vol. 4:235-242, in press.
- Harwood, A. "Guidelines for culturally appropriate care." In A. Harwood (Ed.), *Ethnicity and medical care* (pp. 482-509). Cambridge, MA: Harvard University Press.
- Lenburg, C., Lipson, J.G., Demi, A.S., Blaney, D.R., Stern, P.N., Schultz, P.R., et al. (1995). *Promoting cultural competence in and through nursing education: A critical review and comprehensive plan for action*. Washington, DC: American Academy of Nursing.
- National Multicultural Institute. (1999). "Cultural competency in health care: A guide for trainers." (2nd ed.). Washington, DC: National Multicultural Institute Publications.

Evaluation

- Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Long Grove, IL: Waveland Press, Inc.
- American Evaluation Association (AEA) www.eval.org
- Carmines, E. G., & Zeller, R.A. (1979). *Reliability and validity assessment*. Newbury Park, CA: Sage Publications.
- Converse, J. M., & Presser, S. (1986). *Survey questions : Handcrafting the standardized questionnaire*. Newbury Park, CA: Sage Publications.
- Crocker, L., & Algina, J. (1986). *Introduction to modern and classical test theory*. Fort Worth, TX: Harcourt Brace College Publishers.

- DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Dignan, M. B. (1995). *Measurement and evaluation of health education*. (3rd ed.). Springfield, IL: Charles C Thomas.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1997). *Empowerment evaluation. Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage Publications.
- Green, L. W., & Lewis, F. M. (1986). *Measurement and evaluation in health education and health promotion*. Mountain View, CA: Mayfield.
- Hand, D. J. (2005). *Measurement theory and practice*. London: Arnold Publishers.
- Juvenile Justice Evaluation Center (2001) *Hiring and Working with an Evaluator*. Available at <http://www.jrsa.org/jjec/about/publications/evaluator.pdf>
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage Publications.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Use and interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lorig, K, Stewart, A., Ritter, P., Gonzalez, V., Laurent, D., & Lynch, J. (1996). *Outcome measures for health education and other health education interventions*. Newbury Park, CA: Sage Publishers.
- Miller, D.C. (1991). *Handbook of research design and social measurement*. (5th ed.). Newbury Park, CA: Sage Publications, Inc.
- Morgan, G. A., Leech, N. L., & Barrett, K. C. (2005). *SPSS for introductory and intermediate statistics*. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Netemeyer, R. G., Bearden, W. O, & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: Sage Publications.
- Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory*. New York: Mc Graw Hill.
- Ohio Program Evaluators' Group (OPEG) www.opeg.org
- Patton, M. Q. (2001). *Qualitative research and evaluation methods*. (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage Publications.
- Polit, D. F., & Beck, C. T. (2003). *Nursing research. Principles and methods*. (7th. ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Rogers, H. J., Swaminathan, H., & Hambleton, R. K. (1991). *Fundamentals of item response theory*. Thousand Oaks, CA: Sage Publications.
- Sarvella, P.D., & McDermott, R.J. (1993). *Health education and measurement: A practitioners perspective*. Madison, WI: Brown and Benchmark.
- Shultz, K. S., & Whitney, D. J. (2004). *Measurement theory in action: Case studies and exercises*. Thousand Oaks, CA: Sage Publications.
- Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use*. Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2000). *Using multivariate statistics*. (4th ed.). Boston, MA: Allyn & Bacon.
- Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks, CA: Sage Publications.
- Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2004). *Measurement in nursing and health research*. New York: Springer Publishing Company.
- Wholey, J.S., Hatry, H.P., and Newcomer, K.E. (1994). *Handbook of practical program evaluation*. San Francisco, CA: Jossey-Bass Publishers.

Health Disparities

- Association of State and Territorial Health Officials, National Association of County and City Health Officials, and HRSA (undated) *Health Departments Take Action. A compendium of state and local models addressing racial and ethnic disparities in health*. Retrieved January 31, 2006 from www.astho.org/pubs/nabookfull.pdf
- Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, et al. (2004). *Literacy and health outcomes*. Evidence Report/Technology Assessment No. 87 (Prepared by RTI International—University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication No. 04-E007-2. Rockville, MD: Agency for Healthcare Research and Quality. January 2004.

Institute of Medicine (2002). *Unequal treatment: Confronting racial and ethnic disparities in health care*. Washington, D.C.: National Academy Press.

Randall, V. (2005) *Race, Health Care, and the Law*. Available at <http://academic.udayton.edu/health/index.htm>

United States Department of Health and Human Services. (2000). *Healthy People 2010*. Washington, DC: U.S. Government Printing Office.

Health Literacy

Agency for Healthcare Research and Quality. (2004). *Literacy and health outcomes*. Report Number 04-E007-1.

Bresolin, L. (1999). Health literacy. *The Journal of the American Medical Association (JAMA)*, 281, 552-557.

Davis, T. C., Frederickson, D. D., Arnold, C., Murphy, P W, Herbst, M, Bocchini, J A (1998). A polio immunization pamphlet with increased appeal and simplified language does not improve comprehension to an acceptable level. *Patient Education Counsel*, 33, 5-37.

Institute of Medicine (2004). *Health Literacy: A prescription to end confusion*. Washington, DC: The National Academy Press.

Kalichman, S. C. and Rompa, D. (2000). "Functional Health Literacy is Associated with Health Status and Health-Related Knowledge in People Living with HIV-AIDS," Vol.25(4):337-344.

Michielutte, R. Bahnson, J. Dignan, M.B., Schroeder, E. M. (1992). The use of illustrations and narrative text style to improve readability of a health education brochure. *Journal of Cancer Education*, 7, 251-60.

Murphy, P. W., Chesson, A.L., Walker, L. Arnold, C L, Chesson, L M. (2000). Comparing the effectiveness of video and written material for improving knowledge among sleep disorders clinic patients with limited literacy skills. *Southern Medical Journal*, 93, 297-304.

National Library of Medicine. (2000). *Current bibliographies in medicine: Health literacy*. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.

Ratzan, SC, Parker, RM. (2000). Introduction. In: National Library of Medicine Current Bibliographies in Medicine: Health Literacy. NLM Publication Number CBM 2000-1. Selden CR, Zorn M, Ratzan SC, Parker, RM, Editors. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.

Williams, M.V., Baker, D.W., Honig, E.G., Lee, T.M., Nowlan, A.. (1998). Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest*, 114, 1008-1015.

Williams, M.V., Baker, D.W., Parker, R.M., Nurss, J.R. (1998). Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. *Archives of Internal Medicine*, 158, 166-172.

Multicultural Evaluation

California Endowment (2005). *Commissioning Multicultural Health Evaluation: A Foundation Resource Guide*. Retrieved February 1, 2006 from www.calendow.org/reference/publications/pdf/evaluations/TCE0510-2004_Commissioning_.pdf

California Endowment (2003) *Overview of Multicultural and Culturally Competent Program Evaluation: Issues, Challenges, & Opportunities*. Available at www.calendow.org/reference/publications/pdf/evaluations/TCE0509-2004_Overview_of_Mu.pdf

California Endowment (2003). *Toward Culturally Competent Evaluation in Health and Mental Health*. Available at www.calendow.org/reference/publications/pdf/evaluations/TCE0511-2004_Toward_Cultura.pdf

California Endowment (2003). *Foundation Leaders' Perspectives on Culturally Competent Evaluation*. Available at www.calendow.org/reference/publications/pdf/evaluations/TCE0508-2004_Foundation_Lea.pdf

California Endowment (2003). *Voices from the Field: Health and Evaluation Leaders on Multicultural Evaluation*. Available at www.calendow.org/reference/publications/pdf/evaluations/TCE0515-2005_Voice_From_The.pdf

Georgetown University Center for Child and Human Development (2003). *A Guide to Choosing and Adapting Culturally and Linguistically Competent Health Promotion Materials*. Retrieved February 1, 2006 from gucchd.georgetown.edu/nccc/documents/Materials_Guide.pdf.

U.S. Department of Health and Human Services, Office of Civil Rights (2001) *Policy Guidance Title VI Prohibition Against National Origin Discrimination As It Affects Persons With Limited English Proficiency*. Available at hhs.gov/ocr/lep/guide.html

Reliability and Validity

DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd ed.). Thousand Oaks, CA: Sage Publications.

Dilorio, C. (2005). *Measurement in Health Behavior and Health Education: Methods for Research and Education*. San Francisco: Jossey-Bass.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: Sage Publications.

Schnall, J. G., & Rich, J. (1994). Finding measurement tools. Retrieved October 21, 2005 from <http://healthlinks.washington.edu/howto/measure.html> .

Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use*. New York: Oxford University Press.

University Library at University of Illinois at Chicago (2004). Health instruments. Available at <http://www.uic.edu/depts/lib/lhsu/resources/hlthinst.shtml> .

Selection of an Evaluator

Juvenile Justice Evaluation Center (2001) *Hiring and Working with an Evaluator*. Available at <http://www.jrsa.org/jjec/about/publications/evaluator.pdf>

Stufflebeam, D.L. (1999) Evaluation contracts checklist. Available at <http://www.wmich.edu/evalctr/checklists/contracts.pdf>

X. REFERENCES

- Agency for Healthcare Research and Quality. (2004). *Literacy and health outcomes*. Report Number 04-E007-1.
- Airhihenbuwa, C. O. (1995). *Health and culture: Beyond the Western paradigm*. Thousand Oaks, CA: Sage Publications
- Arnold, C.L., Davis, T.C., Berkel, H.J., Jackson, R.H., Nandy, I. (2001) [Smoking status, reading level, and knowledge of tobacco effects among low-income pregnant women](#). *Preventive Medicine*, Vol. 32:313-20.
- Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.). *Education for the intercultural experience*. Yarmouth, ME: Intercultural Press.
- Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, et al. (2004). *Literacy and health outcomes*. Evidence Report/Technology Assessment No. 87 (Prepared by RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication No. 04-E007-2. Rockville, MD: Agency for Healthcare Research and Quality. January 2004.
- Bouffard, S., & Little, P. M. D. (2004). *Detangling data collection: Methods for gathering data*. Harvard Family Research Project, Harvard Graduate School of Education. Retrieved February 3, 2006 from <http://www.gse.harvard.edu/hfrp/projects/afterschool/resources/snapshot5.html>
- Bresolin, L. (1999). Health literacy. *The Journal of the American Medical Association (JAMA)*,. Vol. 281:552-557.
- Cournoyer, D. E., & Klein, W. C. (2000). *Research methods for social work*. Need page numbers Allyn & Bacon, Boston.
- Davis, T. C., Frederickson, D. D., Arnold, C., Murphy, P W, Herbst, M, Bocchini, J A (1998). A polio immunization pamphlet with increased appeal and simplified language does not improve comprehension to an acceptable level. *Patient Education Counsel*, Vol. 33: 5-37.
- Deason, L. M. & Wallace, S. L. (2005), Does a shared culture between nursing home staff and residents influence quality of life?”, *International Journal of Diversity in Organisations, Communities, and Nations*, Vol. 4:235-242, in press.
- Endo, T., Joh, T., & Yu, H. C. (2003). *Voices from the field: Health and evaluation leaders on multicultural evaluation*. Woodland Hills, CA: The California Endowment.
- Institute of Medicine (2002). *Unequal treatment: Confronting racial and ethnic disparities in health care*. Washington, D.C.: National Academy Press.

- Institute of Medicine (2004). *Health Literacy: A prescription to end confusion*. Washington, DC: The National Academy Press.
- Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage Publications, Inc.
- Lenburg, C., Lipson, J.G., Demi, A.S., Blaney, D.R., Stern, P.N., Schultz, P.R., et al. (1995). *Promoting cultural competence in and through nursing education: A critical review and comprehensive plan for action*. Washington, DC: American Academy of Nursing.
- Michielutte, R. Bahnson, J. Dignan, M.B., Schroeder, E. M. (1992). The use of illustrations and narrative text style to improve readability of a health education brochure. *Journal of Cancer Education*, Vol. 7: 251-60.
- Murphy, P. W., Chesson, A.L., Walker, L. Arnold, C L, Chesson, L M. (2000). Comparing the effectiveness of video and written material for improving knowledge among sleep disorders clinic patients with limited literacy skills. *Southern Medical Journal*, Vol. 93:297-304.
- National Library of Medicine. (2000). *Current bibliographies in medicine: Health literacy*. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.
- Royce, D. (1992). *Program evaluation*. Nelson-Hall, Inc., Chicago.
- Sharma, M, & Deepak, S. (2001). A participatory evaluation of community-based rehabilitation program in North Central Vietnam. *Disability and Rehabilitation*, Vol. 23: 352-358.
- Sharma, M. (2004). Qualitative research methods. In K. E. Berg and R. W. Latin, *Essentials of modern research methods in health, physical education, and recreation*. (2nd ed.) pp. 213-224. Philadelphia: Lippincott.
- Tripp-Reimer, T. (1997). Ethnicity, aging and chronic illness. In L. Swanson and T. Tripp Reimer (Eds.) *Chronic illness and the older adult*. pp. 112-135. New York, NY: Springer Publishing Company.
- United Way of America. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.
- Williams, M.V., Baker, D.W., Honig, E.G., Lee, T.M., Nowlan, A.. (1998). Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest*, Vol. 114:1008-1015.
- Williams, M.V., Baker, D.W., Parker, R.M., Nurss, J.R. (1998). Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. *Archives of Internal Medicine*, Vol. 158:166-172.

APPENDICES

Appendix A. Bennett's Model of Cultural Competence

There are many developmental models of cultural competence in the field. One commonly referenced model was created by Milton Bennett (1993) and consists of 6 stages of moving from ethno-centrism to ethno-relativism. The 6 stages of the Bennett model include:

1. Denial: an individual denies that cultural differences exist.
2. Defense: an individual acknowledges the existence of certain cultural differences, but because those differences are threatening to her/his own reality and sense of self, the individual constructs defenses against those differences.
3. Minimization: an individual acknowledges cultural differences, but trivializes them, believing that human similarities far outweigh the differences.
4. Acceptance: an individual recognizes and values cultural differences without evaluating those differences as positive or negative.
5. Adaptation: an individual develops and improves skills for interacting and communicating with people of other cultures.
6. Integration: an individual not only values a variety of cultures, but are constantly defining their own identity and evaluating behavior and values in contrast to and in concert with a multitude of cultures. Rising above the limitations of living in one cultural context, these individuals integrate aspects of their own original cultural perspective with those of other cultures.

Appendix B. Culturally Competent Methodologies

Mixed Method Approaches. Mixed method approaches are at the heart of most descriptions of multicultural evaluation strategies. Qualitative methods—such as interviews, focus groups, ethnographic studies or observations—are described as critical for capturing a rich level of data on the cultural context of the communities under study. As much as surveys can quantitatively document differences in outcomes, interviews or focus groups can be useful in gathering more nuanced information on *contributing factors* to those differences. This is key especially within diverse populations where issues such as cultural misunderstanding, socio-economic barriers (i.e., transportation, child care, etc.), or even institutionalized racism may otherwise go undetected. Observation and ethnography can be important for capturing the nuances of cross-cultural nonverbal communication (i.e., between teachers and students) to get closer to the understanding of truth. Further, qualitative approaches such as interviews or focus groups were described as simply more appropriate for gathering accurate data on specific populations that may have built a resistance to written surveys. While respondents highlight qualitative data as a method necessary for culturally competent evaluations, none do so to the exclusion of quantitative approaches. Rather, most health and evaluation leaders with whom we spoke conclude that using mixed methods — using both quantitative and qualitative data sources in culturally competent evaluation — helps reign in cultural bias, leads to more thorough research and gives evaluators more evidence to counteract accusations of nonvalidity to any one research method. Mixed method approaches were described by some as especially critical within the health field, where data such as rates of access and utilization of services, health insurance coverage, disease incidence, etc., drive decision-making and funding at the federal level.

[Qualitative methods] afford the evaluator an opportunity to get data that is more rich than a quick and dirty survey...to tease out what would be viable interventions to problems that [target populations] face. We have not paid enough attention to mining the wisdom and insight of these folks. —Ricardo Millett, funder

Multiculturalism to me means the way you put an issue, a problem, a study, a questionnaire in context...and then tools are appropriately applied...knowing when and how to use them is going to be informed by your understanding of the context, how you develop questions, translate the questions, interpret interviews, etc. can all be shaped by the way we understand and contextualize a particular given problem.—Ricardo Millett, funder

Triangulation. Triangulation involves using independent or different sources of measurement that corroborate the same findings and is also offered as a strategy. Triangulation offers similar benefits as using mixed methods in an evaluation study, and has an added benefit of multiple sources of data supporting the same findings to evaluation questions. For example, in evaluating an HIV- prevention program, one evaluator used quantitative HIV/AIDS infection rates data and then complemented them with ethnographic profiles of those going through the prevention program. Another evaluator described how they triangulate data obtained from reports, interviews, written surveys and site visits. They use these four data sources to look for convergence and divergence among the sources.

Reflexive staffing. Several health leaders and evaluators mention recruiting evaluators who reflect the communities studied as another important culturally competent evaluation strategy. Ensuring that evaluation teams are composed of professionals who reflect the community is important for a number of reasons. First, diverse staff often possesses knowledge about the culture, traditions and behavioral patterns of their own ethnic group because they were raised in a

manner similar to other group members. Second, sharing a particular racial or ethnic background helps evaluators to be more sensitive to particular issues because there is a likelihood that they themselves have faced similar situations (e.g., racial discrimination and heightened awareness of power differentials that exist between majority and minority communities). This awareness of cultural differences and nuances are likely to shape how they develop or adapt outcome indicators and valid cross-cultural measurements of program “success” or “failure.” Third, appearance matters. When community members see an evaluator from their ethnic or racial background, they are more likely to think that the person is someone with whom they can identify and trust, and that the evaluation will respond with greater sensitivity to their concerns.

It isn't easy for nonminority evaluators to understand issues of inclusion and power...More often than not, nonminority evaluators are more geared toward the bottom line, getting efficient and effective answers, sacrificing the sloppy and often tough work of understanding context and would more favor taking the shortcut.

While it is clear that hiring diverse evaluation staff can improve the quality of evaluations, at the same time, the evaluation and health leaders urge the adoption of a more sophisticated definition of culturally competent evaluators that look beyond race and ethnicity. As discussed earlier, culturally competent evaluators and culturally appropriate practices are not just a function of race and ethnicity, but are also influenced by a multitude of other factors such as cultural heritage, geographic community and socioeconomic status.

Attention to reflexive staffing is an important initial consideration, but capitalizing on diversity means that all individuals who work with communities of difference need to be prepared to establish bonds of trust across gaps of culture and race.

Engagement of Community Members. A few health leaders and evaluators describe engagement of community members *themselves* as a strategy for gaining cultural insight on a community and ensuring that the design and implementation of an evaluation honors a community’s cultural context. Engaging community members as co-researchers empowers community-based groups and builds their ability to conduct self-evaluation. One evaluation leader gave an example of how inclusion of communities in the evaluation design helped to inform their context. When asking different home care workers across four ethnic groups their conception of mental illness, the study found that each group had a different concept of what it meant to be mentally ill. Because of the four different concepts, defining the outcomes, then, became more complicated and time consuming, yet this added measure led to a better evaluation. Some evaluators describe an approach of always making sure a “cultural translator” is a part of their evaluation team so they are sure they are capturing nuances of a culture that might be different than theirs. A cultural translator is either an evaluator him or herself, a research associate, or a beneficiary stakeholder. For example, one African American evaluator described doing some evaluation with a Native American community, where he employed a Native American research associate on his team. This team member was able to describe the significance of a community baking an earthcake and how this process helped to describe the community mental health of the Navajo tribe. This evaluator also described how another cultural translator trained in survey design was able to point out to him that a Likert scale (a type of survey response scale) does not particularly work well with the Hmong population, who generally see a program as “good” or “bad” and nothing in between.

Appendix C. Outcome Indicators for Health Promotion/Disease Prevention Programs

Note: many of the short-term, intermediate, and long term outcomes were taken from *Healthy People 2010*, the ten year health objectives established by the U.S. Department of Health and Human Services in the year 2000.

CANCER PREVENTION			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes*
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of health care providers receiving intervention to increase screening rates. ▪ Number of hours of educational intervention provided 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ participants showing improvement in knowledge, attitude, and/or skills scores related to: <ul style="list-style-type: none"> ○ cancer screening, ○ healthy dietary changes, ○ physical activity or ○ tobacco cessation. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ People who use skin cancer prevention measures ▪ Women over 18 who have ever: <ul style="list-style-type: none"> ○ received a Pap test. ○ received a Pap test within the past three years. ▪ Women aged 40 years and older who have received a mammogram within the preceding 2 years. ▪ Adults over 50 years who: <ul style="list-style-type: none"> ○ received a fecal occult blood test (FOBT) within the preceding 2 years for colorectal cancer screening. ○ received a sigmoidoscopy for colorectal cancer screening.. ▪ Persons aged 2 years and older who consume: <ul style="list-style-type: none"> ○ at least two daily servings of fruit. ○ at least three daily servings of vegetables, with at least one-third being dark green or orange vegetables. ○ < 10 % of calories from saturated fat. ○ < 30 % of calories from total fat. ▪ Adults who, in the past 12 months, report having had an examination to detect oral and pharyngeal cancers. ▪ Cigarette/spit tobacco/cigar use by adults ▪ Cigarette/spit tobacco/cigar use by adolescents 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ overall or organ specific cancer death rates. ▪ cancer survivors living 5 years or longer after diagnosis.

* Most long term outcomes will be beyond the ability of health promotion programs to effect in a limited project time. They are included in the table to illustrate what short term and intermediate outcomes are meant to eventually achieve.

CANCER PREVENTION			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes*
		in the past month. <ul style="list-style-type: none"> ▪ Smoking cessation attempts by adult smokers. ▪ Tobacco use cessation attempts by adolescent smokers. ▪ Physicians and dentists who counsel their at-risk patients about tobacco use cessation, physical activity, and cancer screening. ▪ Oral and pharyngeal cancers detected at the earliest stage in the targeted populations. 	

CARDIOVASCULAR DISEASE			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of participants receiving physical activity intervention. ▪ Number of educational/other interventions offered. ▪ Number of activities participated in. ▪ Number of hours of direct contact with program participants. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ adolescents who disapprove smoking. ▪ participants showing improvement in knowledge/attitude/skill scores related to: <ul style="list-style-type: none"> ○ early signs of heart disease and heart attack, ○ stress reduction, ○ healthy dietary changes ○ physical activity or ○ tobacco cessation. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ Adults who have had their blood cholesterol checked within the preceding 5 years. ▪ Persons with coronary heart disease who have their LDL-cholesterol level treated to a goal of less than or equal to 100 mg/dL. ▪ Older adults who have participated during the preceding year in at least one organized health promotion activity related to cardiovascular disease. ▪ Persons aged 2 years and older who consume: <ul style="list-style-type: none"> ○ at least two daily servings of fruit. ○ at least three daily servings of vegetables, with at least one-third being dark green or orange vegetables. ○ <10 % of calories from saturated fat. ○ <30 % of calories from total fat. ▪ Adults who: <ul style="list-style-type: none"> ○ engage in no leisure-time physical activity. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ coronary heart disease death rates. ▪ mean total blood cholesterol levels among adults. ▪ adults with high total blood cholesterol levels. ▪ adults who are at a healthy weight. ▪ adults who are obese. ▪ children and adolescents who are obese. ▪ nonsmokers exposed to environmental tobacco smoke as measured by serum cotinine levels.

CARDIOVASCULAR DISEASE			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
		<ul style="list-style-type: none"> ○ engage regularly, preferably daily, in moderate physical activity for at least 30 minutes per day. ○ engage in vigorous physical activity that promotes the development and maintenance of cardiorespiratory fitness 3 or more days per week for 20 or more minutes per occasion. ▪ Adolescents who: <ul style="list-style-type: none"> ○ engage in moderate physical activity for at least 30 minutes on 5 or more of the previous 7 days. ○ engage in vigorous physical activity that promotes cardiorespiratory fitness 3 or more days per week for 20 or more minutes per occasion. ○ adolescents who view television 2 or fewer hours on a school day. ▪ Adolescents and adults who make trips by: <ul style="list-style-type: none"> ○ walking. ○ bicycling. ▪ Cigarette/spit tobacco/cigar use by adults. ▪ Cigarette/spit tobacco/cigar use by adolescents in the past month. ▪ Worksites employing majority of employees from targeted minority population(s) that offer: <ul style="list-style-type: none"> ○ nutrition or weight management classes or counseling to their employees. ○ employer-sponsored physical activity and fitness programs to their employees. 	

DIABETES			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of participants counseled about diet. ▪ Number of educational interventions offered. ▪ Number of hours of direct contact with program participants. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ participants showing improvement in knowledge/attitude/skills scores related to: <ul style="list-style-type: none"> ○ early signs of diabetes detection, ○ diabetes management, ○ healthy dietary changes, or ○ physical activity. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ Adults with diabetes who: <ul style="list-style-type: none"> ○ obtain an annual urinary microalbumin measurement. ○ have a glycosylated hemoglobin measurement at least once a year. ○ have an annual dilated eye examination. ○ have at least an annual foot examination. ○ who have at least an annual dental examination. ○ take aspirin at least 15 times per month. ○ perform self-blood-glucose-monitoring at least once daily. ▪ Physician office visits made by patients with a diagnosis of diabetes that include counseling or education related to diet and nutrition. ▪ Adults who: <ul style="list-style-type: none"> ○ engage in no leisure-time physical activity. ○ engage regularly, preferably daily, in moderate physical activity for at least 30 minutes per day. ○ engage in vigorous physical activity that promotes the development and maintenance of cardiorespiratory fitness 3 or more days per week for 20 or more minutes per occasion. ▪ Adolescents who: <ul style="list-style-type: none"> ○ engage in moderate physical activity for at least 30 minutes on 5 or more of the previous 7 days. ○ engage in vigorous physical activity that promotes cardiorespiratory fitness 3 or more days per week for 20 or more 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ new cases of diabetes per 1,000 population per year. ▪ overall cases of clinically diagnosed diabetes per 1,000 population per year. ▪ adults with diabetes whose condition has been diagnosed. ▪ diabetes death rate among people. ▪ diabetes-related deaths among persons with diabetes. ▪ deaths from cardiovascular disease in persons with diabetes. ▪ pregnant women with gestational diabetes. ▪ frequency of foot ulcers in persons with diabetes. ▪ lower extremity amputations in persons with diabetes. ▪ new cases of end-stage renal disease (ESRD) in diabetics. ▪ kidney failure due to diabetes in diabetics. ▪ adults who are at a healthy weight. ▪ adults who are obese children and adolescents who are obese.

DIABETES			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
		minutes per occasion. <ul style="list-style-type: none"> ▪ Persons with diabetes who receive formal diabetes education. ▪ Persons who have been appropriately counseled about: <ul style="list-style-type: none"> ○ physical activity or exercise. ○ diet and nutrition. ○ smoking cessation. ▪ Pregnant women who receive early and adequate prenatal care. ▪ Worksites employing majority of employees from targeted minority population(s) that offer nutrition or weight management classes or counseling to their employees. 	

HYPERTENSION			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of participants counseled about diet. ▪ Number of participants counseled about medications. ▪ Number of educational interventions offered. ▪ Number of hours of direct contact with program participants. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ participants showing improvement in knowledge/attitude/skills scores related to: <ul style="list-style-type: none"> ○ hypertension. ○ stress reduction. ○ healthy dietary changes. ○ physical activity or ○ tobacco cessation. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ Adults with high blood pressure who are taking action (for example, losing weight, increasing physical activity, or reducing sodium intake) to help control their blood pressure. ▪ Adults who have had their blood pressure measured within the preceding 2 years and can state whether their blood pressure was normal or high. ▪ Persons who have been appropriately counseled about: <ul style="list-style-type: none"> ○ physical activity or exercise. ○ diet and nutrition. ○ smoking cessation. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ adults diagnosed with high blood pressure. ▪ adults with high blood pressure whose blood pressure is under control.

INFANT MORTALITY

Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of participants counseled about diet in pregnancy. ▪ Number/hours of educational interventions offered. ▪ Number of hours of direct contact with program participants. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ participants showing improvement in knowledge/attitude/skills scores related to infant mortality reduction behaviors and activities. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ Pregnant women who attend a series of prepared childbirth classes. ▪ Healthy full-term infants who are put down to sleep on their backs. ▪ Nonpregnant women aged 15 to 44 years who consume of at least 400 µg of folic acid each day from fortified foods or dietary supplements. ▪ Abstinence from alcohol, cigarettes, and illicit drugs among pregnant women. ▪ Mothers who: <ul style="list-style-type: none"> ○ breastfeed their babies in early post partum period. ○ breastfeed their babies at 6 months. ○ breastfeed their babies at one year. ▪ Pregnant women who receive early and adequate prenatal care. ▪ Mothers who achieve a recommended weight gain during their pregnancies. ▪ Newborn bloodspot screening, follow-up testing, and referral to services in the targeted minority populations. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ fetal and infant deaths during perinatal period (28 weeks of gestation to 7 days or more after birth). ▪ all infant deaths (within one year). ▪ neonatal deaths (within the first 28 days of life). ▪ postneonatal deaths (between 28 days of life and one year). ▪ deaths due to Sudden Infant Death Syndrome (SIDS). ▪ maternal complications during hospitalized labor and delivery. ▪ low birth weight (LBW) infants. ▪ very low birth weight (VLBW) infants. ▪ preterm births. ▪ live births at 32 to 36 weeks of gestation. ▪ live births at less than 32 weeks of gestation.

SUBSTANCE ABUSE

Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of educational interventions offered. ▪ Number of hours of direct contact with program participants. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ adolescents who: <ul style="list-style-type: none"> ○ disapprove of substance abuse. ○ perceive great risk associated with consuming five or more alcoholic drinks at a single occasion once or twice a week. ○ perceive great risk associated with smoking marijuana once a month. ○ perceive great risk associated with using cocaine once a month. ○ disapprove of smoking. ▪ participants showing improvement in knowledge/attitude/skills scores related to substance abuse (including alcohol and tobacco) prevention and cessation. 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ Adolescents who report that they rode, during the previous 30 days, with a driver who had been drinking alcohol. ▪ Adolescents not using alcohol or any illicit drugs during the past 30 days. ▪ Adolescents reporting use of marijuana during the past 30 days. ▪ Adults reporting use of any illicit drug during the past 30 days. ▪ High school seniors/college students/adults engaging in binge drinking of alcoholic beverages during the past month. ▪ Adolescents aged 12 to 17 years engaging in binge drinking of alcoholic beverages during the past month. ▪ Steroid use among adolescents. ▪ Adolescents who use inhalants. ▪ Cigarette/spit tobacco/cigar use by adults. ▪ Cigarette use/spit tobacco/cigar use by adolescents in the past month. ▪ Smoking cessation attempts by adult smokers. ▪ Smoking cessation attempts by adolescent smokers. ▪ Abstinence in past month from alcohol/binge drinking/cigarette smoking/illicit drugs by pregnant mothers. ▪ Inmates receiving substance abuse treatment in correctional institutions. ▪ Worksites employing majority of employees from targeted minority population(s) that have formal smoking policies that prohibit smoking or limit it to separately ventilated areas. ▪ Physicians who counsel their at-risk patients 	<p>Outcome Indicators:</p> <ul style="list-style-type: none"> ▪ deaths caused by alcohol-related motor vehicle. ▪ injuries caused by alcohol-related motor vehicle crashes. ▪ deaths caused by drug-related motor vehicle crashes. ▪ injuries caused by drug-related motor vehicle crashes. ▪ cirrhosis deaths. ▪ drug-induced deaths. ▪ average age of first use of alcohol in adolescents age 12-17 year old that remain alcohol free. ▪ average age of first use of marijuana in adolescents age 12-17 year old that remain alcohol free. ▪ adolescents never using alcohol. ▪ adolescents never using illicit drugs.

SUBSTANCE ABUSE			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
		about tobacco use cessation. <ul style="list-style-type: none"> ▪ Dentists who counsel their at-risk patients about tobacco use cessation. ▪ Persons who have been appropriately counseled about smoking cessation. ▪ Persons who have been appropriately counseled about reduced alcohol consumption. 	

VIOLENCE			
Outputs (process outcomes)	Initial Outcomes	Intermediate outcomes	Long-term outcomes
<ul style="list-style-type: none"> ▪ Number of participants receiving educational intervention. ▪ Number of educational interventions offered. ▪ Number of hours of direct contact with program participants. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ participants showing improvement in knowledge/attitudes/skills scores related to violence prevention. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ Physical fighting in past 12 months among children or adolescents. ▪ Weapon carrying by children or adolescents. ▪ Adolescents or children who are counseled about violence prevention. ▪ Adults who are counseled about violence prevention. 	Outcome Indicators: <ul style="list-style-type: none"> ▪ deaths caused by homicides. ▪ child maltreatment incidence. ▪ child maltreatment fatalities. ▪ the rate of physical assault by current or former intimate partners. ▪ rate of rape or attempted rape. ▪ rate of sexual assault other than rape. ▪ rate of physical assault. ▪ intentional injuries resulting from alcohol-related and illicit drug-related violence. ▪ Suicide. ▪ suicide attempts by adolescents. ▪ work-related assaults.

APPENDIX D.

II. GLOSSARY OF SELECTED EVALUATION TERMS

Activities are what a program does with its inputs—the services it provides—to fulfill its mission. Examples include conducting a screening program for breast cancer among Asian American community, educating the African American community members in a neighborhood about how to be physically active, and providing social support for sustaining physical activity in a group of Hispanic Americans. Program *activities* result in *outputs*.

Benchmarks are performance data that are used for comparative purposes. Baseline benchmark data is collected prior to the participants receiving the project activities. A program can use its own data as a baseline benchmark against which to compare future performance. It, also, can use data from another program as a benchmark. In the latter case, the other program often is chosen because it is exemplary and its data are used as a target to strive for, rather than as a baseline.

Community impact is a concept reflecting the collective effect of an intervention beyond the individual-level changes captured via specific outcome indicators. For example, a program seeking to prevent adolescent childbearing in specific neighborhoods, may over time, be able to document declines in the teen pregnancy rate for the targeted geographic areas.

Inputs are resources a program uses to achieve program objectives. Examples include staff, volunteers, facilities, equipment, curricula, and money (funding and/or program budget). A program uses *inputs* to support *activities*.

Instruments are used to obtain information. The term, instrument, is often used interchangeably with the terms tool, survey, measurement, and questionnaire. It is also sometimes referred to as a scale.

Logic Model is a planning tool used to clarify and illustrate what your project intends to do and what it hopes to accomplish and impact. A logic model can be considered a map that you develop to clarify and communicate what your project intends to do and its presumed impact.

Outcomes are benefits for participants during or after their involvement with a program. Outcomes include changes in knowledge, skills, attitudes, values, behaviors, conditions or health status. Examples of outcomes include greater knowledge of nutritional needs, improved reading skills, more effective responses to conflict, getting a job and losing weight.

For a particular program, there can be various "levels" of outcomes, with initial outcomes leading to intermediate outcomes which then lead to longer-term outcomes. For example, a screening program for cervical cancer detection among African American women may show a 20% increase in Pap test for women over 18 years of age during a one year period. This program if sustained over a longer period of time and broadened in its reach would in turn lead to decreased death rates due to cervical cancer.

Outcome indicators are the specific items of information that track a program's success in achieving its outcomes. They describe observable, measurable characteristics or changes that represent achievement of an outcome. For example, a program whose desired outcome is that participants pursue a healthy lifestyle could define "healthy lifestyle" as not smoking; maintaining a recommended weight, blood pressure, and cholesterol level; getting at least two hours of exercise each week; and wearing seat belts consistently. The number and percent of program participants who demonstrate these behaviors then is an *indicator* of how well the program is doing with respect to the outcome. Pre- and post-test surveys, attendance rosters, and numerous measurement tools are used as indicators to track a program's success.

Outcome targets are numerical objectives that specify a program's level of achievement on its outcomes. After a program has had experience with measuring outcomes, it can use its findings to set targets for the number and percent of participants expected to achieve desired outcomes in the next reporting period. It, also, can set targets for the amount of change it expects participants to experience. For example, an outcome target might state that 80% of participants will demonstrate knowledge gain from pre- to post-test survey results.

Outputs are products of a program's activities, such as the number of meals provided, classes taught, brochures distributed, or participants served. Another term for "outputs" is "units of service." A program's *outputs* should produce desired *outcomes* for the program's participants.

Reliability refers to the consistency of an instrument to yield the same results when the instrument is used to measure the same thing time and time again. An instrument is said to be reliable, therefore, when it consistently measures a concept or phenomenon with accuracy. Reliability asks the question: Do repeated applications under similar conditions yield consistent results? A reliable instrument is like a reliable watch – it is not subject to extraneous factors such as the race or weight or height or income of the person wearing it, the month of the year, day of the week, the temperature, etc.

Scale is a cluster of items intended to measure a single concept. For example, in a study designed to measure attitudes about diabetes, an instrument may contain 'knowledge of diabetes' scale, a 'fear of diabetes' scale, and 'action taken to prevent or manage diabetes' scale. An instrument may contain one or more scales depending on the variables that need to be measured in order to efficiently examine the concept being studied.

Validity is the correspondence between what a measuring device is supposed to measure and what it really measures. An instrument is said to be valid when it measures what it was designed/intended to measure.